Sparse Discrete Ordinates Method in Radiative Transfer

Konstantin Grella* and Christoph Schwab

{konstantin.grella, christoph.schwab}@sam.math.ethz.ch

Seminar für Angewandte Mathematik

Eidgenössische Technische Hochschule

CH-8092 Zürich

Switzerland

June 30, 2011

The stationary monochromatic radiative transfer equation (RTE) is a partial differential transport equation stated on a five-dimensional phase space, the Cartesian product of physical and angular domain.

We solve the RTE with a Galerkin FEM in physical space and collocation in angle, corresponding to a discrete ordinates method (DOM). To reduce the complexity of the problem and to avoid the "curse of dimension", we adapt the sparse grid combination technique to the solution space of the RTE and show that we obtain a sparse DOM which uses essentially only as many degrees of freedom as required for a purely spatial transport problem. For smooth solutions, the convergence rates deteriorate only by a logarithmic factor.

We compare the sparse DOM to the standard full DOM and a sparse tensor product approach developed earlier with Galerkin FEM in physical space and a spectral method in angle. Numerical experiments confirm our findings.

Keywords: Radiative transfer, discrete ordinates method, combination technique, sparse grids.

2010 Mathematics Subject Classification: 35Q79, 65N12, 65N30, 65N35.

*Corresponding author. Address: Seminar für Angewandte Mathematik, HG J 59, Rämistrasse 101, 8092 Zürich, Switzerland. Phone: +41 44 632 6362. Fax: +41 44 632 1104.

1 Introduction

1.1 Radiative transfer

In this work we solve the stationary monochromatic radiative transfer problem [see e. g. 15] without scattering defined on a bounded Lipschitz domain $D \subset \mathbb{R}^d$, where $d = 2, 3$.

We would like to find the radiative intensity $u(\mathbf{x}, \mathbf{s})$, $u : D \times S \to \mathbb{R}$, S being the d_S-sphere with $d_S = 1, 2$, that satisfies

$$\mathbf{s} \cdot \boldsymbol{\nabla}_x u(\mathbf{x}, \mathbf{s}) + \kappa(\mathbf{x}) u(\mathbf{x}, \mathbf{s}) = \kappa(\mathbf{x}) I_b(\mathbf{x}), \quad (\mathbf{x}, \mathbf{s}) \in D \times S \tag{1a}$$

$$u(\mathbf{x}, \mathbf{s}) = g(\mathbf{x}, \mathbf{s}), \qquad \mathbf{x} \in \partial D,\ \mathbf{s} \cdot \mathbf{n}(\mathbf{x}) < 0\ . \tag{1b}$$

In this problem statement, $\kappa \geq 0$ is the absorption coefficient, $I_b \geq 0$ the blackbody intensity, $g \geq 0$ the radiation entering the domain or wall emission, and $\mathbf{n}(\mathbf{x})$ the outer unit normal on the boundary. In the presence of media in the domain, the problem formulation could include scattering terms [15]. However, here we neglect scattering since we investigate a new approach to discretize Eq. (1). Furthermore, we assume $g = 0$, i. e. the domain boundaries are non-emissive or "cold".

An introduction to the topic of radiative heat transfer is given by Modest [15]. Apart from Monte Carlo methods, standard solution approaches to the radiative transfer problem are the discrete ordinates method and the method of spherical harmonics. Frank [5] gives an overview of these numerical methods for radiative transfer. State-of-the-art methods and applications are compiled by Kanschat et al. [12].

In the discrete ordinate method (DOM) or S_N-approximation, Eq. (1) is solved for N fixed directions spanning the full range in solid angle. The method is simple to implement and thus popular, but in order to capture very localized features of the solution in the $\mathbf{s}$-dependence a fine angular resolution is usually necessary. Also, the method suffers from so-called *ray effects*, in which the mesh structure of the discretization is reflected in the solution [13].

In the method of spherical harmonics or P_N-approximation, the intensity is expanded into a truncated series of spherical harmonics in solid angle, resulting in a coupled system of PDEs in space. Often used is the P_1-approximation, in which (1a) is reduced to a diffusion equation. In general, though, higher orders lead to a sharp increase in mathematical complexity when boundary conditions are to be satisfied [16]. For smooth solutions, the spherical harmonics method exhibit spectral convergence, which makes them a popular and promising approach for radiative transfer problems where smoothness in the solution is expected when absorption or scattering are present.

The system of partial differential equations arising from the S_N- or P_N-approximation is discretized with finite differences or a finite element method. Manteuffel et al. [14], for instance, solve a least squares formulation with spherical harmonics in solid angle and finite elements in space. Kanschat [11] combines the discrete ordinate method with a stabilized streamline diffusion FEM in the physical domain.

All these methods suffer from the "curse of dimension", the low rate of convergence in terms of number of degrees of freedom due to the high dimensionality of the radiative

transfer problem which is stated in five dimensions for $(d, d_S) = (3, 2)$. The accuracy of the solution does not scale in the same way as the computational complexity so that accurate discretizations quickly become prohibitively expensive.

Widmer et al. [21] have developed a method to overcome the curse of dimension in the context of a wavelet discretization of the angular domain. In their sparse tensor product method, they discretize physical and angular domain with hierarchical and wavelet finite elements, respectively, and then select only the most relevant finite element product combinations to construct the trial space for the solution. Provided that the absorption coefficient $\kappa(\mathbf{x})$ and blackbody intensity $I_b(\mathbf{x})$ are sufficiently smooth, their method achieves a log-linear complexity in the number of degrees of freedom while convergence rates deteriorate only by a logarithmic factor. Their method is also suited for the optically thin regime, i. e. for small κ.

In earlier work [7], we have shown that the sparse tensor product method can also be combined with a spectral discretization involving spherical harmonics, as already suggested by Widmer et al. [22]. The advantages of sparse tensorization then also carry over to a combination of hierarchical finite elements in physical space and spectral discretization in solid angle. Together with additional geometry adapted angular basis functions to satisfy boundary conditions strongly, the sparse tensor spherical harmonics method makes it possible to include spherical harmonics of high order in the solution of the radiative transfer problem without incurring the "curse of dimension".

In this paper we investigate the application of the sparse grid combination technique to the popular discrete ordinates method. Ordinarily, the DOM discretizes the angular space by picking a number of discrete directions. Then a purely physical transport subproblem is solved for each of the fixed directions and the subproblem solutions are combined to form a solution of the RTE. With the sparse grid combination technique [10], or hereafter more generally referred to as the sparse tensor combination technique, not all the subproblems are solved with the same physical resolution. Instead, we compute high resolution solutions only for few directions, while the resolution is lowered for many other directions in a way to achieve our complexity and accuracy goal: provided that the exact solution to the RTE is sufficiently smooth, the convergence rates with respect to the number of degrees of freedom (DoFs) deteriorate only by a logarithmic factor, while the problem size is reduced to a purely physical problem (also up to logarithmic factors).

For the solution of the physical subproblems, we derive a variational formulation from a least squares ansatz and discretize using a Galerkin FEM with hierarchical basis. This allows us to simply add up solutions to different subproblems without interpolation losses.

1.2 Structure and notation of the paper

The paper is organized as follows: In Section 2 we present the discrete ordinates method for the radiative transfer problem (1).

In Section 3 we describe the sparse grid combination technique with a generalization to arbitrary tensor product spaces.

Section 4 explains our sparse DOM and elaborates on complexity and convergence estimates.

Finally, Section 5 underlines the analytical derivations with results from numerical experiments. We also compare the performance to the ordinary full DOM and the sparse/full tensor spherical harmonics approximation developed earlier.

Throughout the present paper, we will use the notation $a \lesssim b$ ($a \simeq b$) if there exists a constant $0 < C < \infty$ with $a \leq Cb$ ($a \leq Cb$ and $a \geq C^{-1}b$). The constants in the inequalities may depend on additional quantities such as angles in a mesh or the dimensions d and d_S.

2 Discrete ordinates method

2.1 Discretization in angle

As a way of discretization of the radiative transfer problem (1), the discrete ordinates method first discretizes the angular domain by picking a discrete set of directions $S_N = \{\mathbf{s}_j\}_{j=1}^{M_S} \subset S$ for which the parametrized RTE is to be solved:

$$\mathbf{s}_j \cdot \boldsymbol{\nabla}_x u(\mathbf{x}, \mathbf{s}_j) + \kappa(\mathbf{x}) u(\mathbf{x}, \mathbf{s}_j) = \kappa(\mathbf{x}) I_b(\mathbf{x}), \quad \mathbf{x} \in D,\ \mathbf{s}_j \in S_N \tag{2a}$$

$$u(\mathbf{x}, \mathbf{s}_j) = g(\mathbf{x}, \mathbf{s}_j), \qquad \mathbf{x} \in \partial D,\ \mathbf{s}_j \cdot \mathbf{n}(\mathbf{x}) < 0\ . \tag{2b}$$

We choose the number of directions to be $M_S = 2N + 3$ for $d_S = 1$ and $M_S = (N+1)^2 + 3$ for $d_S = 2$ so that the parameter N is comparable to the number of harmonic functions up to order N on the sphere. Two or three directions for $d_S = 2$, resp., are added to obtain more than one direction for the case $N = 0$, too.

Effectively, problem (2) requires the solution of M_S purely physical transport problems for fixed directions. If we set $u_j(\mathbf{x}) := u(\mathbf{x}, \mathbf{s}_j)$ and $g_j(\mathbf{x}) := g(\mathbf{x}, \mathbf{s}_j)$, we obtain a system of partial differential equations in the physical domain which are even uncoupled in the case without scattering. We note this system in operator form for compact formulation by introducing $T_j u_j := (\mathbf{s}_j \cdot \boldsymbol{\nabla}_x + \kappa(\mathbf{x}))u_j$ and $f := \kappa(\mathbf{x}) I_b(\mathbf{x})$:

$$T_j u_j = f, \quad j = 1, \dots M_S, \tag{3a}$$

$$u_j|_{\Gamma_-^{(j)}} = g_j, \quad \Gamma_-^{(j)} = \{\mathbf{x} \in \partial D : \mathbf{s}_j \cdot \mathbf{n}(\mathbf{x}) < 0\}. \tag{3b}$$

By $\Gamma_-^{(j)}$ we denote the inflow part of the physical domain boundary which depends on $\mathbf{s}_j$. From here on, we consider the case of $g_j(\mathbf{x}) = 0$ for all j.

2.2 Discretization in physical space

As a second step, we discretize the physical domain with a Galerkin FEM. As the physical problem is a convection-reaction equation of hyperbolic character, the standard Galerkin approach results in an unstable scheme [e. g. 24]. Therefore we stabilize the method by constructing a variational problem based on a least squares functional

[14]:

$$J(u_j) := \frac{1}{2}\left(\epsilon(T_j u_j - f), T_j u_j - f\right)_{L^2}, \tag{4}$$

where

$$\epsilon(\mathbf{x}) = \begin{cases} 1, & \kappa(\mathbf{x}) < \kappa_0, \\ \frac{1}{\kappa(\mathbf{x})}, & \kappa(\mathbf{x}) \geq \kappa_0 \end{cases} \tag{5}$$

with $\kappa_0 \approx 0.134$ [for details, see 21].

In (4), we used a short-hand notation for the inner product:

$$(u, v)_{L^2} := (u, v)_{L^2(D)} = \int_D u\, v \,\mathrm{d}\mathbf{x}. \tag{6}$$

The associated L^2-norm will be denoted by $\|\cdot\|_{L^2(D)} = \|\cdot\|_{L^2}$.

While the solutions u of (1) are in the Hilbert space

$$\mathcal{V}_0 := \{u \in \mathcal{V};\, u = 0 \text{ if } \mathbf{s}\cdot\mathbf{n}(\mathbf{x}) < 0\}, \tag{7}$$

with

$$\mathcal{V} := \{u \in L^2(D \times S) :\ \mathbf{s}\cdot\boldsymbol{\nabla}_x u \in L^2(D \times S)\}, \tag{8}$$

the solutions u_j of (3) must be in the Hilbert space

$$\mathcal{V}^{(j)} := \{u \in L^2(D) :\ \mathbf{s}_j\cdot\boldsymbol{\nabla}_x u \in L^2(D)\} \tag{9}$$

or more precisely, under consideration of the boundary conditions, from the subspace

$$\mathcal{V}_0^{(j)} := \{u \in \mathcal{V}^{(j)};\, u = 0 \text{ on } \Gamma_-^{(j)}\}. \tag{10}$$

The associated norm of $\mathcal{V}^{(j)}$ is $\|\cdot\|_S$, defined by

$$\|u\|_S^2 := \|\mathbf{s}_j\cdot\boldsymbol{\nabla}_x u\|_{L^2(D)}^2 + \|u\|_{L^2(D)}^2. \tag{11}$$

Our bilinear form is based on the transport operator part in the least squares functional (4):

$$a_j(u, v) := (\epsilon T_j u, T_j v)_{L^2(D)}, \quad u, v \in \mathcal{V}^{(j)}. \tag{12}$$

The source functional accordingly contains the part of the source function from (4):

$$l_j(v) := (\epsilon f, T_j v)_{L^2(D)} \tag{13}$$

With these quantities we can state the linear variational problem: For each $j = 1, \ldots, M_S$, find $\tilde{u} \in \mathcal{V}_0^{(j)}$ such that

$$a_j(\tilde{u}, v) = l_j(v) \quad \forall v \in \mathcal{V}_0^{(j)}. \tag{14}$$

Continuity and coercivity of $a_j(\cdot,\cdot)$ can be proven analogously to Proposition 3.3 by Widmer et al. [21] so that together with continuity of the source functional, the well-posedness of (14) follows readily as long as κ remains bounded in the L^∞-norm.

In the formulation of the variational problem above we obtained a different Hilbert space for each direction. In principle, the solution $u_j(\mathbf{x})$ only needs to have a square-integrable directional derivative along the transport direction $\mathbf{s}_j$. However, in the end we are looking for a solution $u(\mathbf{x}, \mathbf{s})$ for every direction $\mathbf{s}$. For the error analysis later, we are therefore going to assume that the solution $u_j(\mathbf{x})$ is in $H^1(D)$. This enforces a stronger regularity across D on the solution than what would be required by $u_j \in \mathcal{V}^{(j)}$, e. g. line discontinuities transported into D from discontinuous boundary data are excluded.

As in the previous work by Widmer et al. [21], we discretize $H^1(D)$ by choosing a hierarchical sequence of spaces V_D^l on dyadically refined meshes $\mathcal{T}_D^l$, $l = 0, \ldots, L$ over the physical domain:

$$V_D^l := S^{p,1}(D, \mathcal{T}_D^l) \subset H^1(D). \tag{15}$$

They consist of piecewise polynomial functions of degree $p \geq 1$ on the mesh $\mathcal{T}_D^l$ which are continuous in D. Altogether, we obtain a sequence of hierarchic finite dimensional subspaces

$$V_D^0 \subset V_D^1 \subset \ldots \subset V_D^L \subset H^1(D),$$

with the dimension $M_D := \dim V_D^L$.

By choosing V_D^L as discrete trial and test space, the directional dependence of this function space would be eliminated again. However, as we also would like to satisfy the boundary conditions in a strong sense, we have to choose our trial and test space to be $V_{D,0}^{L,j} := V_D^L \cap \mathcal{V}_0^{(j)}$. The function spaces which comply with the boundary conditions necessarily depend on the direction $\mathbf{s}_j$ again because the inflow boundary $\Gamma_-^{(j)}$ depends on $\mathbf{s}_j$.

Then we write the discretized variational problem for each direction $\mathbf{s}_j$ as follows: Find $u_{j,L}(\mathbf{x}) \in V_{D,0}^{L,j}$ such that

$$a(u_{j,L}, v_{j,L}) = l(v_{j,L}) \qquad \forall v_{j,L} \in V_{D,0}^{L,j}. \tag{16}$$

Let $\{\alpha_i(\mathbf{x})\}_{i=1}^{M_D}$ be a basis of V_D^L, then a subset $\{\alpha_i(\mathbf{x})\}_{i \in I_j}$ of the basis elements constitutes a basis for $V_{D,0}^{L,j}$. This subset contains only those physical basis functions which are zero on the inflow boundary for direction $\mathbf{s}_j$. For each j, the discretized problem (16) then leads to a linear system of equations in the $|I_j|$ unknowns u_{ij}. The unknowns u_{ij}, $i \notin I_j$, are determined by the boundary conditions.

2.3 Recovery of solution

To recover an approximation $u_{L,N}(\mathbf{x}, \mathbf{s})$ for the solution $u(\mathbf{x}, \mathbf{s})$ of the original RTE (1), we first assemble the approximate solutions $u_{j,L}$ for each direction $\mathbf{s}_j$ as

$$u_{j,L}(\mathbf{x}) = \sum_{i=1}^{M_D} u_{ij} \alpha_i(\mathbf{x}).$$

In angular space we interpolate between the directions $\mathbf{s}_j$. Let $I_S^N : C^0(S) \to \mathbb{P}_N^{d_S}$ be the interpolation operator from the set of continuous functions on the sphere to the

set of spherical polynomials of degree at most N. For an $f \in C^0(S)$ it is given by $I_S^N f(\mathbf{s}_j) = \sum_{j=1}^{M_S} v_j \beta_j(\mathbf{s}_j)$, $\mathbf{s}_j \in S_N$, where $\beta_j \in \mathbb{P}_N^{d_S}$ is a basis for $\mathbb{P}_N^{d_S}$ and the $v_j \in \mathbb{R}$ are interpolation coefficients determined from the linear system

$$\boldsymbol{B}^\top \mathbf{v} = \mathbf{f}, \tag{17}$$

with $B_{jk} = \beta_j(\mathbf{s}_k)$, $\mathbf{v} = (v_1, \ldots, v_{M_S})^\top$, $\mathbf{f} = (f(\mathbf{s}_1), \ldots, f(\mathbf{s}_{M_S}))^\top$ so that in general each v_j depends on the values of f at all directions $\mathbf{s}_k$. For the interpolation to be well-defined, the set of points S_N on the sphere must be unisolvent, i. e. if $\beta_j(\mathbf{s}_k) = 0\ \forall \mathbf{s}_k \in S_N$, then necessarily $\beta_j \equiv 0$.

In our problem, we would like to interpolate the solutions $u_{j,L}$ for different directions $\mathbf{s}_j$ at a point $\mathbf{x}$ in physical space D. The vector of function values to interpolate is therefore $\mathbf{f} = (u_{1,L}(\mathbf{x}), \ldots, u_{M_S,L}(\mathbf{x}))^\top$. Hence, each interpolation coefficient in our case in general depends on $\mathbf{x}$ via all the solutions $u_{j,L}$. The approximate solution to (1) can therefore be written as

$$u_{L,N}(\mathbf{x}, \mathbf{s}) = \sum_{j=1}^{M_S} v_j(\{u_{ij}\alpha_i(\mathbf{x})\}_{i=1,j=1}^{M_D,M_S})\beta_j(\mathbf{s}). \tag{18}$$

The evaluation of this function requires the solution of the linear system (17), which in general is prohibitively expensive at a complexity of $O(M_S^3)$.

Remark 1. Note, however, that on the circle S^1 the fast Fourier transform provides an $O(M_S \log M_S)$ solution to the interpolation problem, and on the sphere S^2 there exist fast algorithms which compute the interpolation polynomial in $O(M_S^2(\log M_S)^2)$ [17]. As usually mostly derived quantities, which include the intensity u in integrals over the angle, are of interest in applications, we disregard the full solution and calculate e. g. the net emission, incident radiation, or the heat flux as results of our computations. In high-temperature simulations, for instance, the radiative intensity enters the energy equation only as the divergence of the heat flux [15, ch. 1].

3 Sparse tensor combination technique

The sparse grid combination technique on a regular tensor product mesh with mesh-width h as developed by Griebel et al. [8] combines the solutions of $O(\log(h^{-1}))$ different full grids to obtain a solution equivalent to the one on a sparse grid. The theoretical advantages of sparse grid solutions also come into effect for the combination technique: for sufficiently smooth solutions, the number of degrees of freedom reduces from $O(h^{-2})$ to $O(h^{-1}\log(h^{-1}))$ in 2D, which is essentially the complexity of a one-dimensional problem. At the same time, the accuracy deteriorates only by a logarithmic factor from $O(h^2)$ to $O(h^2 \log(h^{-1}))$.

The practical advantages of the combination technique are twofold: First, only full grid problems have to be solved for which standard full grid solvers can be applied. No programming of sparse grid solution approaches is required. Second, as all the single full grid problems are independent, the technique can be trivially parallelized:

one assigns each subproblem to a computing node. During the solution procedure for the full subproblems no communication between the nodes is necessary, only in the end the solutions have to be collected and combined.

Because of these benefits, the sparse grid combination technique seems worth investigating. In order to apply it to our problem, we adapt and generalize the description of the sparse grid combination technique in the following.

3.1 Derivation of the technique

We assume that the domain of the problem is a Cartesian product domain $D_1 \times D_2$, on top of which we approximate the function space $\mathcal{V}_0$, e. g. $L^2(D_1 \times D_2)$, by the tensor product of discrete spaces $V_1^{L_1} \otimes V_2^{L_2}$. For our purposes, we assume that the discrete spaces consist of hierarchic families of spaces $V_1^{l_1}$, $V_2^{l_2}$ with $V_i^0 \subset V_i^1 \subset \ldots \subset V_i^{L_i}$, $i = 1, 2$. Detail spaces $W_i^{l_i}$ denote the increment between $V_i^{l_i-1}$ and $V_i^{l_i}$:

$$V_i^{l_i} = V_i^{l_i-1} \oplus W_i^{l_i}.$$

The approximate full tensor solution can be represented by contributions from detail spaces:

$$u_{L_1,L_2} = \sum_{l_1=0}^{L_1} \sum_{l_2=0}^{L_2} Q_{l_1,l_2} u, \tag{19}$$

where $Q_{l_1,l_2} : \mathcal{V}_0 \to W_1^{l_1} \otimes W_2^{l_2}$ is a projection operator on the tensor product of detail spaces. Then an approximate sparse solution can be given by restricting the range of indices l_1, l_2:

$$\hat{u}_{L_1,L_2} = \sum_{0 \le f(l_1,l_2) \le L_1 L_2} Q_{l_1,l_2} u, \tag{20}$$

with a cutoff profile $f : [0, L_1] \times [0, L_2] \to \mathbb{R}$.

As the sparse solution is a linear combination of the contributions from the detail spaces, it can just as well be assembled from addition and subtraction of different full solutions of smaller index range. As a simple example, we take a look at the case $L_1 = L_2 = L$ with a cutoff $f(l_1, l_2) = Ll_1 + Ll_2$. Then a sparse solution can be expressed by the *combination formula*

$$\tilde{u}_{L,L} = \sum_{i=0}^{L} \sum_{l_1=0}^{L-i} \sum_{l_2=0}^{i} u_{l_1,l_2} - \sum_{i=0}^{L-1} \sum_{l_1=0}^{L-1-i} \sum_{l_2=0}^{i} u_{l_1,l_2}. \tag{21}$$

Instead of solving one sparse problem, one solves $2L + 1$ full problems of smaller size and adds the solutions. One should note, however, that the combined solution $\tilde{u}_{L,L}$ is in general not the same as the original sparse solution $\hat{u}_{L,L}$. It is equivalent in its convergence and complexity properties, though [8]. Only if the projectors onto $V_1^{L_1}, V_2^{L_2}$ commute, the combination technique solution is exactly the sparse solution [10].

The same approach can be used for more general cutoffs f, Fig. 1 illustrates this case. Adaptive sparse grids leading to generalized combination techniques have been

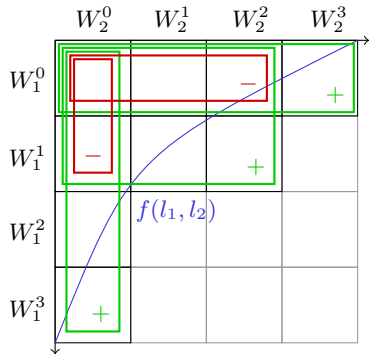

Figure 1: A sparse tensor product space determined by a cutoff profile f is represented by sums and differences of full tensor product spaces of smaller size. Each detail tensor product space $W_1^{l_1} \otimes W_2^{l_2}$ occurs only once in the sparse tensor product space. See Eq. (22).

investigated by Hegland [9] and Garcke [6], for instance. Let $l_2^{\max}$ be the index l_2 obtained by solving $f(l_1, l_2) = L_1 L_2$ w. r. t. l_2 and rounding the result to the next smallest integer, then a sparse solution can be combined like

$$\tilde{u}_{L_1,L_2} = \sum_{i_1 \in \mathcal{I}_a} \left(\sum_{l_1=0}^{i_1} \sum_{l_2=0}^{l_2^{\max}(i_1)} u_{l_1,l_2} - \sum_{l_1=0}^{i_1} \sum_{l_2=0}^{l_2^{\max}(i_1+1)} u_{l_1,l_2} \right). \tag{22}$$

$\mathcal{I}_a$ contains the indices $i_1 \in [0, L_1]$ with unique $l_2^{\max}$, as long as $l_2^{\max}(i_1) \geq l_2^{\max}(L_1)$ (assumed that $l_2^{\max}(\cdot)$ is monotonically decreasing). If several indices have the same $l_2^{\max}$, only the largest index is contained in $\mathcal{I}_a$. In the sum above, we adopted the convention that $l_2^{\max}(i_1) = -1$ for $i_1 > L_1$ so that in this case, the second double sum is empty.

3.2 Complexity estimates

For the derivation of complexity estimates, we assume that the dimension of the component spaces $V_i^{L_i}$ grows exponentially with respect to the levels l_i to a base b_i, and we define the dimension or component complexity as

$$M_i := \dim(V_i^{L_i}) = b_i^{L_i}, \quad i = 1, 2. \tag{23}$$

A full problem solution without sparsification of the solution spaces would therefore entail a complexity of

$$M_{(L_1,L_2)} = M_1 M_2 = b_1^{L_1} b_2^{L_2} \tag{24}$$

Each of the full subproblems of reduced resolution, which we solve in the solution process of the the sparse tensor combination technique, therefore has a complexity of

$M_{(l_1,l_2)}$, if l_1 and l_2 are the levels of resolution of the respective subproblem. The overall complexity $\tilde{M}_{(L_1,L_2)}$ of the sparse tensor combination technique approach for the simple case of a linear cutoff and $L_1 = L_2 = L$ is given in the following lemma.

Lemma 3.1. *Given the dimension of the component spaces $M_i := \dim(V_i^{L_i}) = b_i^{L_i}$, $b_i > 1$, $L_i > 1$, $i = 1, 2$, and $l_2^{\max} = L - l_1$, the complexity of the solution of all subproblems calculated in the sparse tensor combination technique is*

$$\tilde{M}_{(L,L)} \lesssim L^{\theta} \max\{b_1, b_2\}^L, \tag{25}$$

where $\theta = 1$ if $b_1 = b_2$ and zero otherwise.

Proof. In essence, this lemma is identical to one shown by Bieri [2, Lemma 4.2.1]. By summing the contributions of the complexities of the full subproblems we obtain

$$\tilde{M}_{(L,L)} = \sum_{l_1=0}^{L} b_1^{l_1} b_2^{l_2^{\max}} = \sum_{l_1=0}^{L} b_1^{l_1} b_2^{L-l_1} = \sum_{l_1=0}^{L} b_2^L \left(\frac{b_1}{b_2}\right)^{l_1}.$$

If now $b_1 = b_2$, we get $\tilde{M}_{(L,L)} = (L+1)b_2^L$. Otherwise, we assume w. l. o. g. $b_2 > b_1$ and obtain

$$\tilde{M}_{(L,L)} = b_2^L \sum_{l_1=0}^{L} \left(\frac{b_1}{b_2}\right)^{l_1} \simeq \frac{b_1}{b_2} b_1^L - b_2^L \lesssim \max\{b_1^L, b_2^L\}.$$

These cases can be combined into (25). □

4 Sparse discrete ordinates method

In the following, we are going to apply the sparse tensor combination technique to the discrete ordinates method for the radiative transfer equation.

4.1 Derivation of the method

To begin with, we define the L^2 projection operator onto the discrete physical trial space V_D^l as $P_D^l : \mathcal{V}_0 \to V_D^l$, with $P_D^{-1} = 0$ by convention. Also, we recall the definition of the angular interpolation operator $I_S^l : C^0(S) \to \mathbb{P}_l^{d_S}$. Based on the interpolation operator, we can also define a difference operator $\Delta_S^l : C^0(S) \to \mathbb{P}_l^{d_S}$ by

$$\Delta_S^0 := I_S^0, \quad \Delta_S^l := I_S^l - I_S^{l-1},\ l = 1, 2, \ldots \tag{26}$$

In operator notation, the approximation $u_{L,N}$, obtained by the DOM, to the solution u of the RTE (1), as already stated in (18), can hence be expressed as

$$u_{L,N} = \sum_{l_D=0}^{L} \sum_{l_S=0}^{N} \Delta_S^{l_S}((P_D^{l_D} - P_D^{l_D-1})u). \tag{27}$$

To sparsify this full solution, we again limit the range of indices by introducing a summation condition:

$$\hat{u}_{L,N} = \sum_{0 \le f(l_D, l_S) \le LN} \Delta_S^{l_S}((P_D^{l_D} - P_D^{l_D-1})u). \tag{28}$$

With the implicit function $l_S^{\max}(l_D)$, which yields the maximum feasible integer l_S for a value of l_D so that $f(l_D, l_S) \le LN$, the sparse solution may also be written as

$$\hat{u}_{L,N} = \sum_{l_D=0}^{L} \sum_{l_S=0}^{l_S^{\max}(l_D)} \Delta_S^{l_S}((P_D^{l_D} - P_D^{l_D-1})u). \tag{29}$$

In order to calculate the effect of the difference operator $\Delta_S^{l_S}$, we have to calculate the effects of the interpolation operators $I_S^{l_S}$ and $I_S^{l_S-1}$. However, when the sum over l_S is computed, all the interpolated solutions cancel except for the one with highest resolution:

$$\hat{u}_{L,N} = \sum_{l_D=0}^{L} I_S^{l_S^{\max}(l_D)}((P_D^{l_D} - P_D^{l_D-1})u). \tag{30}$$

Now we could solve the $L+1$ full subproblems in the sum in which the degrees of freedom in $W_D^{l_D}$ are determined for a number M_S of directions. In angle we would then interpolate to extend the solution continuously between the directions $\mathbf{s}_j$.

However, if we proceed like this, we lose information about the interrelations of the degrees of freedom from different detail spaces. Instead of solving on the detail spaces separately, we could solve subproblems on $V_D^{l_D}$ and subtract the solution on $V_D^{l_D-1}$:

$$\hat{u}_{L,N} = \sum_{l_D=0}^{L} I_S^{l_S^{\max}(l_D)}(P_D^{l_D}u - P_D^{l_D-1}u). \tag{31}$$

More efficiently, we should only solve and add a subproblem if it is not subtracted again later. This means we do not solve all problems for $l_D = 0, \ldots, L$, but only for the indices l_D in the list

$$\begin{aligned}\mathcal{I}_a &= \{l_D^{(1)}, l_D^{(2)}, \ldots, l_D^{(n_a)}\} \\ &= \{l_D \in [0, L] : \ \forall l'_D \in [0, L] \setminus \{l_D\} : (l_S^{\max}(l'_D) \neq l_S^{\max}(l_D) \vee \\ &\quad (l_S^{\max}(l'_D) = l_S^{\max}(l_D) \ \wedge \ l_D > l'_D))\}.\end{aligned} \tag{32}$$

We also assume the indices of the list to be ordered $l_D^{(1)} < l_D^{(2)} < \ldots < l_D^{(n_a)}$. In words, the list $\mathcal{I}_a$ contains for each value of $l_S^{\max}(l_D) \in [0, N]$ the largest physical level index l_D yielding the respective $l_S^{\max}(l_D)$. Effectively, we obtain the sparse tensor combination technique for the discrete ordinates method:

$$\tilde{u}_{L,N} = \sum_{j=1}^{n_a} \left(I_S^{l_S^{\max}(l_D^{(j)})} P_D^{l_D^{(j)}} u - I_S^{l_S^{\max}(l_D^{(j)})} P_D^{l_D^{(j-1)}} u \right), \quad l_D^{(j)} \in \mathcal{I}_a. \tag{33}$$

We add the convention $l_D^{(0)} = -1$ so that for $j = 1$ there is no subproblem being subtracted. In the last formulation, we again sum over full subproblem solutions with resolution levels $(L, N) = (l_D^{(j)}, l_S^{\max}(l_D^{(j)}))$ with $l_D^{(j)}$ in index set $\mathcal{I}_a$, and for each subproblem except for the first, subtract the solution to the smaller (physically less finely resolved) subproblem with levels $(L, N) = (l_D^{(j-1)}, l_S^{\max}(l_D^{(j)}))$.

4.2 Complexity estimate

In an earlier article, we showed that the complexity of the sparse tensor Galerkin and spectral approach is $\hat{M}_{(L,N)} = O(2^{dL} + L^\theta N^{d_S}) = O(M_D + \log(M_D)^\theta M_S)$, where $\theta = 1$ if $N^{d_S} \simeq 2^{dL}$ [7, Lemma 3.2]. The sparse tensor combination technique DOM exhibits the same complexity estimate, as we will see in the following. Preceding factors will differ from those of ST, though.

Lemma 4.1. *Assumed that the number of angular directions is $M_S = |\mathcal{S}_N| \simeq N^{d_S}$, the dimension of the physical component space is $M_D := \dim(V_D^L) \simeq 2^{dL}$, and $l_2^{\max} = 2^{\lfloor \log_2(N+1) \rfloor / L(L-l_D)}$, the complexity of the solution of all subproblems calculated in the sparse tensor combination technique is*

$$\tilde{M}_{(L,N)} \lesssim L^\theta \max\{2^{dL}, N^{d_S}\}, \tag{34}$$

where $\theta = 1$ if $N^{d_S} \simeq 2^{dL}$ and zero otherwise.

Proof. We reduce the case to the situation in Lemma 3.1. We sum all subproblem complexities with maximum levels $l_S^{\max}$ for all $0 \le l_D \le L$:

$$\tilde{M}_{(L,N)} \lesssim \sum_{l_D=0}^{L} M_{(l_D, l_S^{\max}(l_D))} \simeq \sum_{l_D=0}^{L} (l_S^{\max})^{d_S} 2^{dL} \tag{35}$$

$$= \sum_{l_D=0}^{L} (\underbrace{2^{d_S \lfloor \log_2(N+1) \rfloor / L}}_{=:b_S})^{(L-l_D)} (\underbrace{2^d}_{=:b_D})^{l_D} \tag{36}$$

From here on, we apply Lemma 3.1 and obtain eventually

$$\tilde{M}_{(L,N)} \lesssim \begin{cases} (L+1) b_S^L & \text{if } b_S = b_D, \\ \max\{b_D, b_S\}^L & \text{else} \end{cases} \simeq \begin{cases} L N^{d_S} & \text{if } N^{d_S} \simeq 2^{dL}, \\ \max\{2^{dL}, N^{d_S}\} & \text{else.} \end{cases} \tag{37}$$

In terms of component complexities, we have

$$\tilde{M}_{(L,N)} \lesssim \log(M_D)^\theta \max\{M_D, M_S\}, \tag{38}$$

with $\theta = 1$ if $M_D = M_S$ and zero otherwise as before. □

4.3 Convergence properties

In this section we take a look at the approximation properties of the sparse DOM and derive error estimates for the sparse and full DOM approach with Galerkin FEM in physical space.

Error estimate in physical and angular domain. In physical space, we know that for functions $v \in H^{s+1}(D)$, $s \in [0,\ p]$, projected onto V_D^l from (15) by the projection operator P_D^l, the following approximation properties hold for $l \in \mathbb{N}_0$ (see e. g. Nguyen [18, Lemma 2.3.1], or Braess [3]):

$$\|v - P_D^l v\|_{H^1(D)} \lesssim 2^{-ls} \|v\|_{H^{s+1}(D)},\ s \in [0,\ p]. \tag{39}$$

Furthermore we require an estimate for the interpolation error on the angular domain. To this end, we recall the angular interpolation operator $I_S^N : C^0(S) \to \mathbb{P}_N^{d_S}$, a mapping to the space $\mathbb{P}_N^{d_S}$ of spherical polynomials. $\mathbb{P}_N^{d_S}$ contains all polynomials restricted to the d_S-sphere S of degree less or equal to N. We can therefore write for the interpolation error of a function $w \in C^0(S)$:

$$\|w - I_S^N w\|_{C^0(S)} = \|(w - p) - I_S^N (w - p)\|_{C^0(S)}, \tag{40}$$

with $p \in \mathbb{P}_N^{d_S}$ an arbitrary polynomial. From there, we arrive at the quasi-optimality condition

$$\|w - I_S^N w\|_{C^0(S)} \lesssim (1 + \|I_S^N\|) \inf_{p \in \mathbb{P}_N^{d_S}} \|w - p\|_{C^0(S)}, \tag{41}$$

in which the operator norm

$$\|I_S^N\| = \|I_S^N\|_{C^0(S) \to C^0(S)} = \sup_{w \in C^0(S) \setminus \{0\}} \frac{\|I_S^N w\|_{C^0(S)}}{\|w\|_{C^0(S)}} \tag{42}$$

occurs as the Lebesgue constant determining the stability of the interpolation operator I_S^N in angular space.

Despite the fact that polynomial interpolation on the sphere has been studied extensively in the past, the precise behavior of the Lebesgue constant in terms of N remains unknown for general d_S. A theoretical lower bound is given by the optimal growth rate for all linear projections, the growth of the $C^0(S) \to C^0(S)$ operator norm of the L^2-orthogonal projection as a map from $C^0(S)$ to $\mathbb{P}_N^{d_S}$, which is $O(N^{(d_S-1)/2})$ [20].

The same authors also quote Reimer [19] on the theoretical upper bound $\|I_S^N\| \lesssim M_S \lesssim N^{d_S}$ for the interpolation operator, which holds as long as the system of interpolation points maximizes $|\det(\boldsymbol{B})|$, where $\boldsymbol{B}$ is the system matrix from (17). For the case $d_S = 2$, this estimate reads as $\|I_S^N\| \lesssim (N+1)^2$.

Also for $d_S = 2$, another estimate for an upper bound due to Reimer [19] is $(N + 1)(\lambda_{\text{avg}}/\lambda_{\min})^{1/2}$, with λ_{avg} and $\lambda_{\min}$ being the average and minimum, resp., eigenvalues of the matrix of the reproducing kernel $G_N(\mathbf{s}, \mathbf{s}') = \sum_{n=0}^{N} \sum_{m=1}^{m_{d_S,N}} Y_{nm}^{d_S}(\mathbf{s}) Y_{nm}^{d_S}(\mathbf{s}')$ evaluated for all combinations of two interpolation points $\mathbf{s}, \mathbf{s}' \in S_N$. In this, $Y_{nm}^{d_S}$ are the spherical interpolation polynomials. However, apart from $\lambda_{\text{avg}}/\lambda_{\min} \geq 1$ for $d_S \geq 2$ and $N \geq 3$, the dependence of the eigenvalue ratio on N and d_S is unclear.

Empirically obtained growth rates of the Lebesgue constant fortunately indicate that many extremal point systems have better stability properties than the proved bound of $(N+1)^2$ on the 2-sphere. Womersley and Sloan [23] obtain points that yield

$\|I_S^N\| \approx 0.7N + 1.8$ by directly minimizing the norm of the interpolation operator numerically. Different point systems can be generated from the notion of so-called spherical designs which possess good qualities for quadrature as well as interpolation, as the growth of the Lebesgue constant for these systems also seems to be close to linear with N. An et al. [1] find $\|I_S^N\| \approx 0.8025(N+1)^{1.12}$ for their well-conditioned spherical designs.

Remark 2. For $d_S = 1$, which is the setting of our numerical experiments later on, the situation is much clearer. It is known that equidistant points on the circle are optimal for trigonometric interpolation as they yield the smallest Lebesgue constant, and that the Lebesgue constant grows like $O(\log(N))$ [4, p. 554]. Equidistant points also lead to exponential convergence for quadrature of periodic integrands if the trapezoidal rule is used for integration.

To summarize and to continue estimating the convergence properties of the method, we will assume

$$\|I_S^N\| \lesssim N^{d_S - \alpha} \tag{43}$$

in the following, where $\alpha \in \mathbb{R}$ may still depend on the dimension d_S and the set of interpolation points. For $d_S = 2$, we expect $\alpha \approx 1$ or slightly smaller, based on the reported results above.

If moreover $w \in H^t(S)$, $t > 1$, we can estimate the best N-term approximation by

$$\inf_{p \in \mathbb{P}_N^{d_S}} \|w - p\|_{L^2(S)} \lesssim N^{-t} \|w\|_{H^t(S)}, \tag{44}$$

see e. g. Grella and Schwab [7, Lemma 3.3]. In total, we obtain the estimate for the interpolation error in angular space as

$$\|w - I_S^N w\|_{L^2(S)} \lesssim N^{-t+(d_S-\alpha)} \|w\|_{H^t(S)}. \tag{45}$$

Error estimate for full DOM. In this method, we solve a physical transport problem with Galerkin FEM of M_D degrees of freedom for each of the M_S transport directions. To recover an approximation of the original solution, we would interpolate between the directions. Below, we estimate the error of this approximative solution in the full method. However, in practice we are more interested in computing quantities depending on integrals over the angle because these occur in energy balance equations of simulations and because we can avoid the cost of interpolation.

Theorem 4.2. *The full discrete ordinates approximation $u_{L,N} = I_S^N P_D^L u$ of a function $u \in H^{s+1,t}(D \times S)$, $s \in [0, p]$, $t \in \mathbb{N}_0$, satisfies the asymptotic error estimate*

$$\|u - I_S^N P_D^L u\|_{H^{1,0}(D\times S)} \lesssim \max\{2^{-sL}, N^{-t+(d_S-\alpha)}\} \|u\|_{(H^{1+s,t})(D\times S)}. \tag{46}$$

Proof. We begin by splitting the error in two parts and summing the contributions

from the detail spaces (the domain $D \times S$ is omitted from the norms for brevity):

$$\begin{aligned}\|u - I_S^N P_D^L u\|_{H^{1,0}} \leq \| \sum_{l_D=0}^{L} \sum_{l_S=N+1}^{\infty} \Delta_S^{l_S}(P_D^{l_D} - P_D^{l_D-1})u\|_{H^{1,0}} + \\ \| \sum_{l_D=L+1}^{\infty} \sum_{l_S=0}^{\infty} \Delta_S^{l_S}(P_D^{l_D} - P_D^{l_D-1})u\|_{H^{1,0}} \\ \leq \sum_{l_D=0}^{L} \|(\mathrm{Id}_S - I_S^N)(\mathrm{Id}_D - P_D^{l_D})u\|_{H^{1,0}} + \\ \sum_{l_D=0}^{L} \|(\mathrm{Id}_S - I_S^N)(\mathrm{Id}_D - P_D^{l_D-1})u\|_{H^{1,0}} + \\ \| \mathrm{Id}_S(\mathrm{Id}_D - P_D^L)u\|_{H^{1,0}}\end{aligned}$$

We drop the middle term as it does not result in leading order terms in the estimate, and use the estimates on physical domain (39) and angular domain (45):

$$\begin{aligned}\|u - I_S^N P_D^L u\|_{H^{1,0}} &\lesssim N^{-t+(d_S-\alpha)}(2^s + 2^{-sL})\|u\|_{H^{s+1,t}} + 2^{-sL}\|u\|_{H^{s+1,0}} \\ &\lesssim \max\{2^{-sL}, N^{-t+(d_S-\alpha)}\}\|u\|_{H^{s+1,t}},\end{aligned}$$

which is the statement of the theorem. □

Error estimate for sparse DOM. Here we split the problem of resolution levels L and N into subproblems of coarser resolution, solve them and combine these solutions to a sparse approximation of the original RTE solution.

Theorem 4.3. *Given a sparse cutoff f resulting in $l_2^{\max} = 2^{\lfloor \log_2(N+1) \rfloor / L(L-l_D)}$, the sparse discrete ordinates approximation $\tilde{u}_{L,N}$ of a function $u \in H^{s+1,t}(D \times S)$, $s \in [0, p]$, $t \in \mathbb{N}_0$, satisfies the asymptotic error estimate*

$$\|u - \tilde{u}_{L,N}\|_{H^{1,0}(D\times S)} \lesssim L \max\{2^{-sL}, N^{-t+(d_S-\alpha)}\}\|u\|_{H^{1+s,t}(D\times S)}. \tag{47}$$

Proof. We begin the same way as for the full DOM estimate, only the summation

indices now run within different limits:

$$\begin{aligned}\|u-\tilde{u}_{L,N}\|_{H^{1,0}} \leq \| \sum_{l_D=0}^{L} \sum_{l_S=l_S^{\max}+1}^{\infty} \Delta_S^{l_S}(P_D^{l_D}-P_D^{l_D-1})u\|_{H^{1,0}} + \\ \| \sum_{l_D=L+1}^{\infty} \sum_{l_S=0}^{\infty} \Delta_S^{l_S}(P_D^{l_D}-P_D^{l_D-1})u\|_{H^{1,0}} \\ \leq \sum_{l_D=0}^{L} \|(\mathrm{Id}_S - I_S^{l_S^{\max}})(\mathrm{Id}_D - P_D^{l_D})u\|_{H^{1,0}} + \\ \sum_{l_D=0}^{L} \|(\mathrm{Id}_S - I_S^{l_S^{\max}})(\mathrm{Id}_D - P_D^{l_D-1})u\|_{H^{1,0}} + \\ \| \mathrm{Id}_S(\mathrm{Id}_D - P_D^{L})u\|_{H^{1,0}}\end{aligned}$$

We drop the middle term again, use (39) and (45), and insert $l_S^{\max}$:

$$\begin{aligned}\|u-\tilde{u}_{L,N}\|_{H^{1,0}} &\lesssim \sum_{l_D=0}^{L} (l_S^{\max})^{-t+(d_S-\alpha)} 2^{-sl_D})\|u\|_{H^{s+1,t}} + 2^{-sL}\|u\|_{H^{s+1,0}} \\ &= \sum_{l_D=0}^{L} (2^{\lfloor \log_2(N+1)\rfloor / L(L-l_D)})^{-t+(d_S-\alpha)} 2^{-sl_D})\|u\|_{H^{s+1,t}} + \\ &\quad 2^{-sL}\|u\|_{H^{s+1,0}}.\end{aligned}$$

The sum can be estimated by its largest summand and we get:

$$\begin{aligned}\|u-\tilde{u}_{L,N}\|_{H^{1,0}} &\lesssim (L+1)\max\{N^{-t+(d_S-\alpha)}, 2^{-sL}\}\|u\|_{H^{s+1,t}} + \\ &\quad 2^{-sL}\|u\|_{H^{s+1,0}} \\ &\lesssim L\max\{N^{-t+(d_S-\alpha)}, 2^{-sL}\}\|u\|_{H^{s+1,t}},\end{aligned}$$

which concludes this proof. □

5 Numerical experiments

We conduct a number of numerical experiments for $(d, d_S) = (2, 1)$ to compare the performance of the sparse combination (SC) discrete ordinates method to a full collocation (FC) method in which we solve the physical subproblem for each discrete direction with the highest physical resolution. Additionally, we also compare to a sparse tensor (ST) spherical harmonics approximation and its full tensor (FT) version developed earlier [7]. In these last two approaches we use a Galerkin least squares FEM in physical space and a spectral method in angle.

5.1 Algorithm

Our method has been implemented in an algorithm in MATLAB which has not been optimized for performance yet.

First, we calculate the physical stiffness matrices and load vector entries by quadrature of the integrals over the physical domain that arise out of the bilinear form (12). For the quadrature in D, we use a Gauss-Legendre rule in 2D which integrates the terms of the bilinear form involving product combinations of linear physical basis functions exactly up to rounding errors if the absorption coefficient $\kappa(\mathbf{x})$ is constant. In the load vector integration, we iteratively increase the number of quadrature nodes to achieve a relative error tolerance of 10^{-13}.

The directions $\mathbf{s}_j$ are equispaced in the angular domain and chosen in a way to avoid collisions with boundaries of coherent inflow or outflow regions of the faces of the physical domain. Integration in angle is performed by the trapezoidal rule.

For each direction $\mathbf{s}_j$, the linear system to determine the coefficients of the physical basis functions is solved by a diagonally preconditioned conjugate gradient method. We terminate the CG method if the ℓ^2-norm of the dof (degrees of freedom) residual vector is less than 10^{-16}.

5.2 Quantities of interest

As the radiative intensity is a function of several variables, we are going to inspect derived quantities of reduced dimensionality to simplify visualization and error analysis. Such quantities are the incident radiation $G(\mathbf{x})$, the heat flux $\mathbf{q}(\mathbf{x})$, and the net emission $\boldsymbol{\nabla} \cdot \mathbf{q}(\mathbf{x})$, which are defined and related by

$$G(\mathbf{x}) = \int_S u(\mathbf{x}, \mathbf{s}) \mathrm{d}\mathbf{s} \tag{48}$$

$$\mathbf{q}(\mathbf{x}) = \int_S u(\mathbf{x}, \mathbf{s}) \mathbf{s} \mathrm{d}\mathbf{s} \tag{49}$$

$$\boldsymbol{\nabla} \cdot \mathbf{q}(\mathbf{x}) = \kappa(\mathbf{x}) \left(\begin{cases} 4\pi & \text{if } d_S = 2 \\ 2\pi & \text{if } d_S = 1 \end{cases} \right\} I_b(\mathbf{x}) - G(\mathbf{x})). \tag{50}$$

In simulations of high-temperature situations, the radiative intensity enters the calculations in the energy equation as the divergence of the heat flux [15, ch. 1], hence the interest in the net emission.

In experiments without known solution, we compute a reference solution with the discrete ordinates (DO) method and use this solution to estimate the error in the incident radiation $G_{L,N}(\mathbf{x})$ of the numerical solution $u_{L,N}(\mathbf{x}, \mathbf{s})$. In the DO method, the angular domain is discretized into 256 directions, along which we calculate the solution by the method of lines and a standard non-stiff integrator in MATLAB. The line integrals are then interpolated to the FEM mesh in the physical domain corresponding to a resolution of $L = 7$. We compute the relative error in the incident radiation as

$$err(G_{L,N})_X = \|G - G_{L,N}\|_X / \|G\|_X,$$

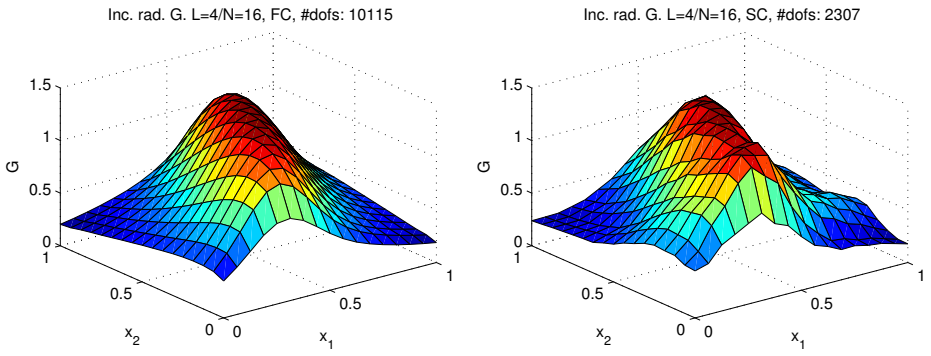

Figure 2: Experiment 1: Incident radiation (left FC, right SC) for $L = 4/N = 16$.

where G is the reference solution of the incident radiation and X stands for the $L^2(D)$- or $H^1(D)$-norm, respectively.

5.3 Experiments

All experiments have been conducted on the physical domain $D = [0, 1]^2$, the unit square, with zero inflow boundary conditions. The absorption coefficient function is constant $\kappa(\mathbf{x}) = 1$. In the collocation methods SC and FC, we use hierarchical hat functions on a square mesh with mesh size $h = 2^{-L}$ as basis in the physical domain. In angular space, we choose $2N + 3$ directions $\mathbf{s}_j$ to discretize the domain so that in the sparse collocation method, each physical subproblem is solved in at least three different directions.

In order to compare the convergence rates to the ones of the sparse tensor spherical harmonics approximation, we use some of the same experiments that have been conducted before [7]. To isolate the convergence rates over the domains D and S, we refine in physical resolution only by incrementing L by 1 and fixing the angular order or vice versa, then N is doubled in each refinement step and L is constant. However, in normal operation one would rather use an equilibration relation to increase the resolution in D and S in a combined manner. Experiments 2 and 3 are examples for combined refinement.

In the SC method, the reported numbers of degrees of freedom represent the sum of all the degrees of freedom of the solved full physical subproblems.

5.3.1 Experiment 1

This example is an application with a degenerate Gaussian on the right hand side:

$$I_b(\mathbf{x}) = \exp\left(-8(\mathbf{x} - \begin{pmatrix}0.5\\0.5\end{pmatrix})^\top \begin{pmatrix}4 & -2\\-2 & 1\end{pmatrix} (\mathbf{x} - \begin{pmatrix}0.5\\0.5\end{pmatrix})\right).$$

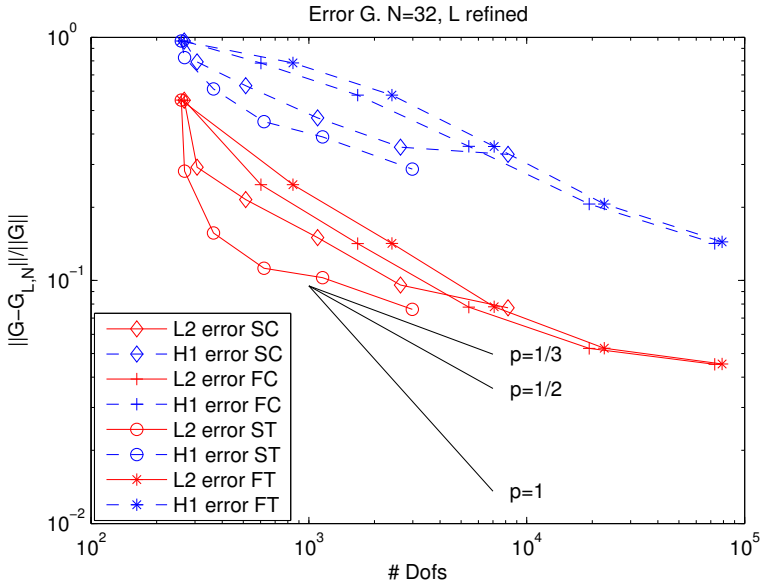

Figure 3: Experiment 1: Convergence in incident radiation (Eq. 48) of numerical solution $G_{L,N}$ to reference solution G for $L = 0, 1, \ldots, 5$ and $N = 32$.

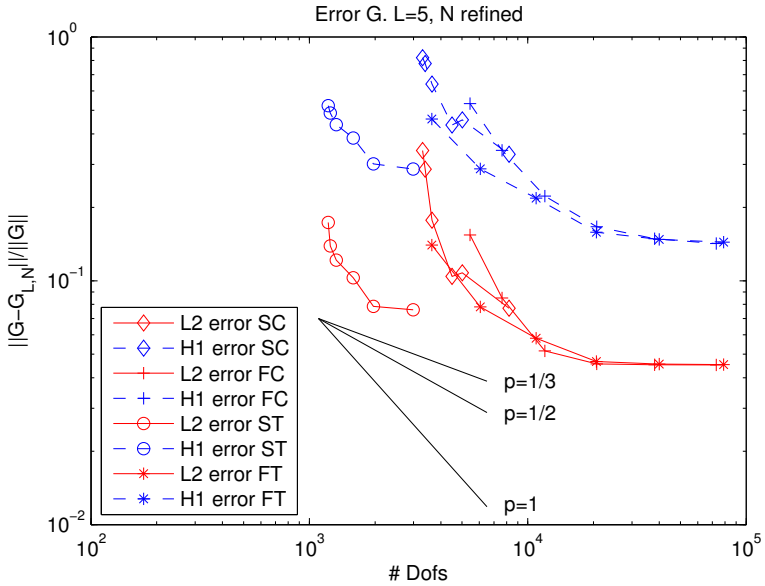

Figure 4: Experiment 1: Convergence in incident radiation (Eq. 48) of numerical solution $G_{L,N}$ to reference solution G for $N = 1, 2, 4, 8, 16, 32$ and $L = 5$.

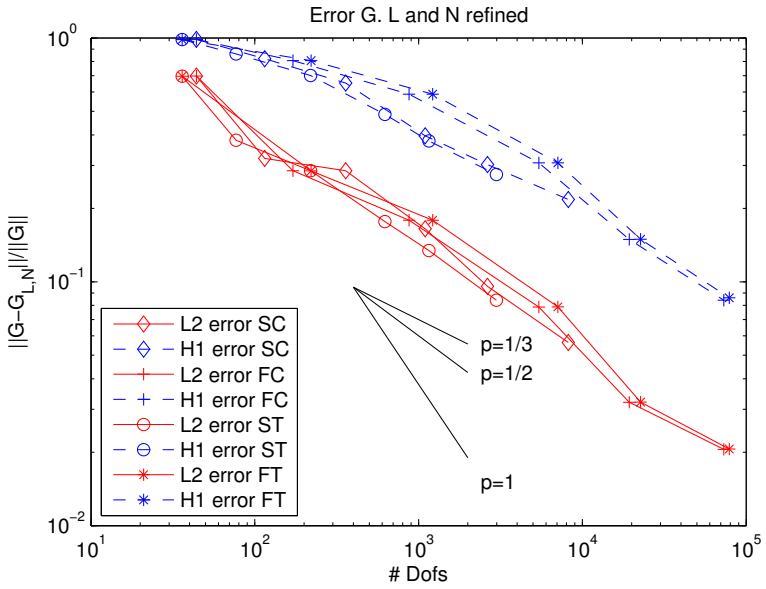

Figure 5: Experiment 2: Error in incident radiation G for combined refinement $L = 0, 1, \ldots, 5$ and $N = \min\{2^{L+2}, 32\}$.

In Fig. 2 we see a good overall qualitative agreement between the FC and SC solutions for the incident radiation for $L = 4$ and $N = 16$. Of course, the SC solution is noticeably rougher than the FC solution, but we have to keep in mind that it has been computed with less than a quarter of the dofs of the FC solution.

Fig. 3 shows that the estimated convergence rates of $1/2$ for the H^1-error in refinement of the physical resolution are actually achieved for all four methods. Best in terms of efficiency, i. e. error per employed degrees of freedom, is the ST method, followed by the SC method. Both full methods perform very similar with a larger error per dofs ratio.

For refinement in angular space, the convergence rates start out very high, but decrease quickly (cf. Fig. 4), even to zero for the full methods. This may be due to a saturation of the error from the physical contribution. The resolution in angle is large enough to accurately resolve the exact solution, but the error from the physical domain remains the same as L is not increased. This is in line with our error estimates as stated in Thms. 4.2 and 4.3: the convergence rate is always determined by the slower rate of physical and angular domain.

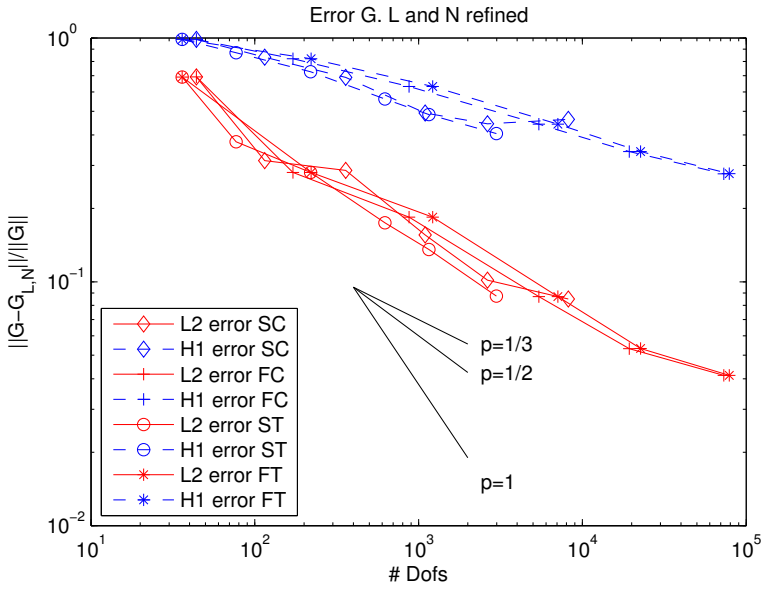

Figure 6: Experiment 3: Error in incident radiation G for combined refinement $L = 0, 1, \ldots, 5$ and $N = \min\{2^{L+2}, 32\}$.

5.3.2 Experiment 2

We use a compactly supported $C^\infty(D)$ bump function on the right hand side:

$$I_b(\mathbf{x}) = \begin{cases} 10^4 \exp\left(\frac{-1}{0.25-(x_1-0.5)^2} + \frac{-1}{0.25-(x_2-0.5)^2}\right) & \text{if } 0.25 < x_1, x_2 < 0.75, \\ 0 & \text{else.} \end{cases}$$

In this experiment, we refine in a combined manner: L is increased from 0 to 5 and $N = \min\{2^{L+2}, 32\}$.

With a smooth right hand side, we observe in Fig. 5 that the expected convergence rate of 1/2 for the H^1-error is again achieved for the full methods FC and FT. The FC method is slightly more efficient with a lower error per employed dofs ratio. The sparse methods SC and ST almost attain a rate of 1/2 in the H^1 error, the difference to 1/2 can be attributed to the log-factor in our estimate (47) stated in Thm. 4.3. In terms of efficiency, both sparse methods are approximately equal, but outperform the full methods. The H^1-error of the SC method is up to 20% lower than the one of the FC method, for the Galerkin FEM/spectral methods ST and FT, the reduction is even slightly greater.

5.3.3 Experiment 3

In this experiment, the blackbody intensity is the characteristic function of a circle in the physical domain:

$$I_b(\mathbf{x}) = \begin{cases} 1 & \text{if } (x_1 - 0.5)^2 + (x_2 - 0.5)^2 < \frac{1}{4^2}, \\ 0 & \text{else.} \end{cases}$$

We again perform combined refinement with $L = 0, \ldots, 5$ and $N = \min\{2^{L+2}, 32\}$.

Due to the lower smoothness of the right hand side, the approximation rates of all methods and in all errors are lower than in the previous experiment. We achieve a rate of about $1/3$ or slightly less in the H^1-error for both full and sparse methods. Again, for the same number of dofs, the ST method yields the smallest error, followed closely by the SC method. With a gap, the FC and FT methods follow. While the convergence of the SC method runs more irregularly, for the finest resolution, the H^1-error even increases slightly, it is still noteworthy that also in this setting of lower smoothness, the sparse methods maintain an advantage in efficiency over the full methods.

6 Conclusion

We investigated a sparsification of the popular discrete ordinates method for radiative transfer in order to alleviate the curse of dimension for this problem. We have shown that the convergence rate in the $H^{1,0}$-error for the sparse tensor combination technique discrete ordinates method is the same as the rate for the full discrete ordinates method ($1/2$ for $d = 2$) up to logarithmic factors, while asymptotically the number of employed degrees of freedom grows only as fast as in a purely physical problem, again up to logarithmic factors.

Numerical experiments have indicated that we are capable of solving the same radiative transfer problems as with a sparse/full Galerkin FE and spectral method with similar approximation properties and that we achieve the estimated convergence rates as long as the solutions are sufficiently smooth.

While the sparse Galerkin FE and spectral method gave an even slightly larger benefit in terms of error per employed degrees of freedom over its full tensor version than the sparse discrete ordinates method, the discrete ordinates method has a big advantage in another respect: as only independent full physical subproblems are solved and combined, implementation of the sparse combination technique can rely on standard full solvers for the physical subproblems. Furthermore, straightforward parallelization via the ordinates opens the door to the solution of large scale applications with HPC.

Acknowledgments. Partial support by the Swiss National Science Foundation (SNF) under project no. 121892 and by ERC AdG no. 247277 STAHDPDE is gratefully acknowledged.

References

[1] C. An, X. Chen, I. H. Sloan, and R. S. Womersley. Well conditioned spherical designs for integration and interpolation on the two-sphere. *SIAM J. Numerical Analysis*, 48:2135–2157, 2010. doi: 10.1137/100795140.

[2] Marcel Bieri. *Sparse tensor discretizations of elliptic PDEs with random input data*. PhD thesis, ETH Zürich, 2009. URL `http://e-collection.ethbib.ethz.ch/view/eth:339`. No. 18598.

[3] Dietrich Braess. *Finite elements: Theory, fast solvers, and applications in elasticity theory*. Cambridge University Press, 3rd edition, 2007.

[4] V. K. Dzyadyk, S. Yu. Dzyadyk, and A. S. Prypik. Asymptotic behavior of Lebesgue constants in trigonometric interpolation. *Ukrainian Mathematical Journal*, 33:553–559, 1981. ISSN 0041-5995. doi: 10.1007/BF01085428.

[5] Martin Frank. Approximate models for radiative transfer. *Bull. Inst. Math. Acad. Sinica (New Series)*, 2:409–432, 2007.

[6] Jochen Garcke. A dimension adaptive sparse grid combination technique for machine learning. In Wayne Read, Jay W. Larson, and A. J. Roberts, editors, *Proceedings of the 13th Biennial Computational Techniques and Applications Conference, CTAC-2006*, volume 48 of *ANZIAM J.*, pages C725–C740, 2007.

[7] Konstantin Grella and Christoph Schwab. Sparse tensor spherical harmonics approximation in radiative transfer. Technical Report 2010-33, SAM, ETH Zürich, October 2010. URL `http://www.sam.math.ethz.ch/reports/2010/33`.

[8] M. Griebel, M. Schneider, and C. Zenger. *Iterative Methods in Linear Algebra*, chapter A combination technique for the solution of sparse grid problems, page 263–281. North-Holland, Amsterdam, 1992.

[9] Markus Hegland. Adaptive sparse grids. In K. Burrage and Roger B. Sidje, editors, *Proc. of 10th Computational Techniques and Applications Conference CTAC-2001*, volume 44 of *ANZIAM J.*, pages C335–C353, 2003.

[10] Markus Hegland, Jochen Garcke, and Vivien Challis. The combination technique and some generalisations. *Linear Algebra and its Applications*, 420(2-3):249–275, 2007. ISSN 0024-3795. doi: 10.1016/j.laa.2006.07.014.

[11] Guido Kanschat. *Parallel and adaptive Galerkin methods for radiative transfer problems*. PhD thesis, University of Heidelberg, 1996. URL `http://archiv.ub.uni-heidelberg.de/volltextserver/volltexte/2006/6331/`.

[12] Guido Kanschat, Wilhelm von Waldenfels, Christian Y. Cardall, Stephen Wright, Simon Arridge, Martin Schweiger, Erik Meinköhn, Maarten Baes, A. Hujeirat, Eugene H. Avrett, Jiří Kubát, Daniela Korčáková, Wolfgang Kalkofen, Dimitris Stamatellos, Anthony P. Whitworth, Avery Meiksin, Sabine Richling, Dmitrij

Semionov, Vladas Vansevičius, and Oliver Dorn. *Numerical Methods in Multidimensional Radiative Transfer*. Springer, 2008. doi: 10.1007/978-3-540-85369-5.

[13] K. D. Lathrop. Ray effects in discrete ordinates equations. *Nucl. Sci. Eng.*, 32 (3):357, 1968.

[14] Thomas A. Manteuffel, Klaus J. Ressel, and Gerhard Starke. A boundary functional for the least-squares finite-element solution of neutron transport problems. *SIAM Journal on Numerical Analysis*, 37(2):556–586, 2000. doi: 10.1137/S0036142998344706.

[15] Michael F. Modest. *Radiative Heat Transfer*. Elsevier, 2nd edition, 2003.

[16] Michael F. Modest and Jun Yang. Elliptic PDE formulation and boundary conditions of the spherical harmonics method of arbitrary order for general three-dimensional geometries. *Journal of Quantitative Spectroscopy & Radiative Transfer*, 109:1641–1666, 2008. doi: 10.1016/j.jqsrt.2007.12.018.

[17] Martin J. Mohlenkamp. A fast transform for spherical harmonics. *Journal of Fourier Analysis and Applications*, 5:159–184, 1999. ISSN 1069-5869. doi: 10.1007/BF01261607. 10.1007/BF01261607.

[18] Hoang Nguyen. *Finite element wavelets for solving partial differential equations*. PhD thesis, University of Utrecht, 2005. URL `http://igitur-archive.library.uu.nl/dissertations/2005-0504-200310/index.htm`.

[19] M. Reimer. *Constructive theory of multivariate functions*. BI Wissenschaftsverlag, Mannheim/Wien/Zürich, 1990.

[20] Ian H. Sloan and Robert S. Womersley. Constructive polynomial approximation on the sphere. *Journal of Approximation Theory*, 103:91–118, 2000. doi: 10.1006/jath.1999.3426.

[21] G. Widmer, R. Hiptmair, and Ch. Schwab. Sparse adaptive finite elements for radiative transfer. Technical Report 01, SAM, ETH Zürich, January 2007. URL `http://www.sam.math.ethz.ch/reports/2007/01`.

[22] G. Widmer, R. Hiptmair, and Ch. Schwab. Sparse adaptive finite elements for radiative transfer. *Journal of Computational Physics*, 227:6071–6105, 2008. doi: 10.1016/j.jcp.2008.02.025.

[23] Robert S. Womersley and Ian H. Sloan. How good can polynomial interpolation on the sphere be? *Advances in Computational Mathematics*, 14(3):195–226, 2001. doi: 10.1023/A:1016630227163.

[24] Matías Ávila, Ramon Codina, and Javier Principe. Spatial approximation of the radiation transport equation using a subgrid-scale finite element method. *Computer Methods in Applied Mechanics and Engineering*, 200:425–438, 2011. doi: 10.1016/j.cma.2010.11.003.